

Soil moisture drought in Europe

Manning, Colin; Widmann, Martin; Bevacqua, Emanuele; Van Loon, Anne; Maraun, Douglas; Vrac, Mathieu

DOI:

[10.1175/JHM-D-18-0017.1](https://doi.org/10.1175/JHM-D-18-0017.1)

License:

Other (please specify with Rights Statement)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Manning, C, Widmann, M, Bevacqua, E, Van Loon, A, Maraun, D & Vrac, M 2018, 'Soil moisture drought in Europe: a compound event of precipitation and potential evapotranspiration on multiple timescales', *Journal of Hydrometeorology*. <https://doi.org/10.1175/JHM-D-18-0017.1>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

Checked for eligibility: 07/08/2018

© Copyright 2018 AMS (For details see: <https://www.ametsoc.org/ams/index.cfm/publications/ethical-guidelines-and-ams-policies/ams-copyright-policy/>)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Soil Moisture Drought in Europe: A Compound Event of Precipitation and Potential Evapotranspiration on Multiple Time Scales

COLIN MANNING AND MARTIN WIDMANN

School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

EMANUELE BEVACQUA

Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

ANNE F. VAN LOON

Department of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

DOUGLAS MARAUN

Wegener Center for Climate and Global Change, University of Graz, Graz, Austria

MATHIEU VRAC

Laboratoire des Sciences du Climat et de l'Environnement (LSCE-IPSL), Centre d'Etudes de Saclay, Gif-sur-Yvette, France


(Manuscript received 27 January 2018, in final form 26 April 2018)

ABSTRACT

Compound events are extreme impacts that depend on multiple variables that need not be extreme themselves. In this study, we analyze soil moisture drought as a compound event of precipitation and potential evapotranspiration (PET) on multiple time scales related to both meteorological drought and heat waves in wet, transitional, and dry climates in Europe during summer. Drought indices that incorporate PET to account for the effect of temperature on drought conditions are sensitive to global warming. However, as evapotranspiration (ET) is moisture limited in dry climates, the use of such drought indices has often been criticized. We therefore assess the relevance of the contributions of both precipitation and PET to the estimation of soil moisture drought. Applying a statistical model based on pair copula constructions to data from FluxNet sites in Europe, we find at all sites that precipitation exerts the main control over soil moisture drought. At wet sites PET is additionally required to explain the onset, severity, and persistence of drought events over different time scales. At dry sites, where ET is moisture limited in summer, PET does not improve the estimation of soil moisture. In dry climates, increases in drought severity measured by indices incorporating PET may therefore not indicate further drying of soil but the increased availability of energy that can contribute to other environmental hazards such as heat waves and wildfires. We therefore highlight that drought indices including PET should be interpreted within the context of the climate and season in which they are applied in order to maximize their value.

1. Introduction

Soil moisture plays a critical role in agriculture and the variability of temperature (Seneviratne et al. 2006).

 Denotes content that is immediately available upon publication as open access.

Corresponding author: Colin Manning, cjm317@bham.ac.uk

As soil moisture observations are sparse, soil moisture drought must be monitored and quantified using indirect methods. These include land surface models that provide physically based estimates of soil moisture (Mitchell et al. 2004; Sheffield et al. 2014) and drought indices that are used as a proxy of soil moisture (Dai et al. 2004; Vicente-Serrano et al. 2012). The simplicity of drought indices is advantageous, but it also leaves their output open to interpretation, which we assess in this article.

DOI: 10.1175/JHM-D-18-0017.1

© 2018 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

Many studies have highlighted the multivariate nature of soil moisture drought as well as the importance of incorporating temperature in drought analysis (Seneviratne et al. 2012b; Teuling et al. 2013; AghaKouchak et al. 2014). Through the inclusion of temperature via potential evapotranspiration (PET) in drought indices such as the standardized precipitation evapotranspiration index (SPEI; Vicente-Serrano et al. 2010), the Palmer drought severity index (PDSI; Palmer 1965), and the reconnaissance drought index (RDI; Tsakiris and Vangelis 2005), studies have been able to analyze how drought conditions may change in a warming climate at regional and global scales (Dai et al. 2004; Dai 2011; Sheffield et al. 2012; Dai 2013; Trenberth et al. 2014; Törnros and Menzel 2014; Vicente-Serrano et al. 2014; Zarch et al. 2015; Stagge et al. 2017).

With increasing temperatures, drought events are expected to set in quicker (Trenberth et al. 2014) and become more severe based on indices incorporating PET (Törnros and Menzel 2014; Zarch et al. 2015; Vicente-Serrano et al. 2014). However, the meaning of this increase in severity for soil moisture according to these indices can be quite unclear due to the differing contribution of PET to soil moisture drought in moisture-limited and energy-limited climates (Seneviratne et al. 2010). Understanding these differences can help in the interpretation of future changes depicted by drought indices and the potential implications they bring at a regional level where impacts of drought are felt.

Soil moisture drought refers to moisture deficits in the upper layer of soil known as the root zone. Soil moisture in the root zone is primarily controlled by antecedent precipitation while excesses in evapotranspiration (ET), related to high temperatures, are required to explain the severity of a negative soil moisture anomaly (Teuling et al. 2013; Seneviratne et al. 2012b). The contribution of ET to soil moisture drought depends on the availability of moisture in the soil for ET to take place (Seneviratne et al. 2010). PET measures the evaporative demand of the atmosphere and indicates the amount of ET that would occur given an unlimited water supply. Under moisture-limited conditions, values of PET and ET can diverge where ET may verge to zero while PET can continue to rise with an increase in temperature (Seneviratne et al. 2010). In such dry conditions, PET and temperature can therefore have little contribution to the estimation of soil moisture (Luo et al. 2017) and lead to drying biases in terms of moisture levels in soil when incorporated into drought indices (Sheffield et al. 2012; Seneviratne 2012).

Describing soil moisture with drought indices requires one to account for antecedent meteorological conditions that soil moisture holds memory of. This is done using integrations of a climatic water balance (precipitation minus PET) varying in length from 1 to 24 months

(e.g., SPEI), or through the use of recursive models (e.g., PDSI). The selection of this integration length for indices such as the SPEI is important; a length that is too short will not capture drought persistence while longer periods can include redundant information (Törnros and Menzel 2014). Studies using the SPEI or RDI to represent soil moisture generally use integration periods between 3 and 6 months (Hirschi et al. 2011; Törnros and Menzel 2014). The PDSI is calculated with monthly integrations and it can hold memory of the previous winter and spring in summer months (Dai et al. 2004).

The use of a climatic water balance implies that PET influences soil moisture over the same time scale as precipitation. However, drying of soil occurs on a daily time scale where excesses in ET can be driven by days of extreme temperature that are filtered out through the use of longer integration periods. Such a feature of long integrations of the climatic water balance can lead to an inability to capture both future changes in drying that may cause droughts to set in quicker in a warmer climate and the occurrence of flash droughts associated with short periods of warm temperature and rapidly decreasing soil moisture (Mo and Lettenmaier 2016).

High temperatures driving excesses in ET can be partly attributed to land–atmosphere interactions induced by deficits in precipitation. By leading to dry soil conditions, low antecedent precipitation is associated with an increased probability of hot days (Hirschi et al. 2011; Mueller and Seneviratne 2012; Whan et al. 2015; Ford and Quiring 2014), amplified extreme temperatures, and the persistence of heat waves (Miralles et al. 2014; Lorenz et al. 2010) that, in turn, can further deplete soil moisture where moisture is available.

The individual roles of precipitation and PET, and that of their dependence driven through land–atmosphere interactions, highlight the compound nature of soil moisture drought. Compound events are a class of events receiving an increased amount of attention in recent times. They encompass a broad range of impacts whose risk is influenced by the dependence between their drivers (Wahl et al. 2015; Hillier et al. 2015; Martius et al. 2016; Bevacqua et al. 2017). Understanding the dependence between hot and dry conditions and their impacts is of great importance. Overlooking nonlinear dependence between hot and dry conditions and crop yields leads to an underestimation of risk in reduced crop yields (Zscheischler et al. 2017), while the bivariate risk of hot and dry summers is underestimated when treating them independently (Zscheischler and Seneviratne 2017). Underlining this importance are findings of an increase in the concurrence of drought and heat wave events (Mazdiyasn and AghaKouchak 2015). Such an increase brings a potential rise in the risk of associated impacts, as the impact arising

TABLE 1. Summary of FluxNet sites used throughout this study.

Site	Site name	Lat	Lon	Site type
a	Dripsey, Ireland	51.99°N	8.75°W	Grassland
b	Hainich, Germany	51.08°N	10.45°E	Forest
c	Klingenberg, Germany	50.89°N	13.52°E	Grassland
d	Oensingen, Switzerland	47.28° N	7.73°E	Grassland
e	Pang/Lambourne, United Kingdom	51.45° N	1.27°W	Forest
f	Le Bray, France	44.72°N	0.77°W	Forest
g	Amplero, Italy	41.9°N	13.6°W	Grassland
h	Las Majadas del Tietar, Spain	39.94°N	5.77°W	Savanna/Grassland
i	Bugacpuszta, Hungary	46.69°N	19.6°E	Grassland
j	Mitra IV Tojal, Portugal	38.48°N	8.02°W	Grassland
k	Vall d'Alinya, Spain	42.15°N	1.45°E	Grassland

from the combination of drought and heat wave events can be greater than the sum of the impacts arising from individual events (Hegerl et al. 2011; Zscheischler et al. 2014).

In this study, we analyze soil moisture drought as a compound event of meteorological drought and heat waves in Europe. We use a conceptual framework developed in Bevacqua et al. (2017) and apply it to a system in which we describe soil moisture as a function of precipitation integrated over preceding months and PET integrated over recent days. This conceptual framework allows us to capture days of extreme temperature within the PET variable and its dependence on antecedent conditions. The framework is implemented via a multivariate statistical model based on pair copula constructions (PCCs; Aas et al. 2009). Copula-based methods, in particular PCCs, provide much flexibility in modeling multidimensional systems (Aas et al. 2009; Bevacqua et al. 2017), including the representation of nonlinear dependence associated with hot and dry conditions (Stage et al. 2015; Zscheischler and Seneviratne 2017). Their use has therefore become quite prominent in the analysis of compound events (Serinaldi et al. 2009; AghaKouchak et al. 2014; Bevacqua et al. 2017; Zscheischler et al. 2017; Zscheischler and Seneviratne 2017). More details on copula, PCCs, and the advantages of their use can be found in the following sections.

We assess the compound nature of soil moisture drought in Europe during the summer months June–August (JJA) at locations in wet, transitional, and dry climates. We aim to demonstrate the individual contributions of precipitation and PET to the estimation of soil moisture drought and highlight where, when, and over what integration period lengths PET and its dependence with precipitation are important for the estimation of soil moisture in a statistical setting. In doing so, we aim to characterize the compound nature of soil moisture drought in differing climates during summer to provide information that may aid with the interpretation

of drought indices incorporating PET and allow further insight to be gained from such indices.

The paper is organized as follows: the data employed in this study as well as the statistical methods involved are described in section 2, the main results are presented in section 3, and a summary and conclusions are provided in section 4.

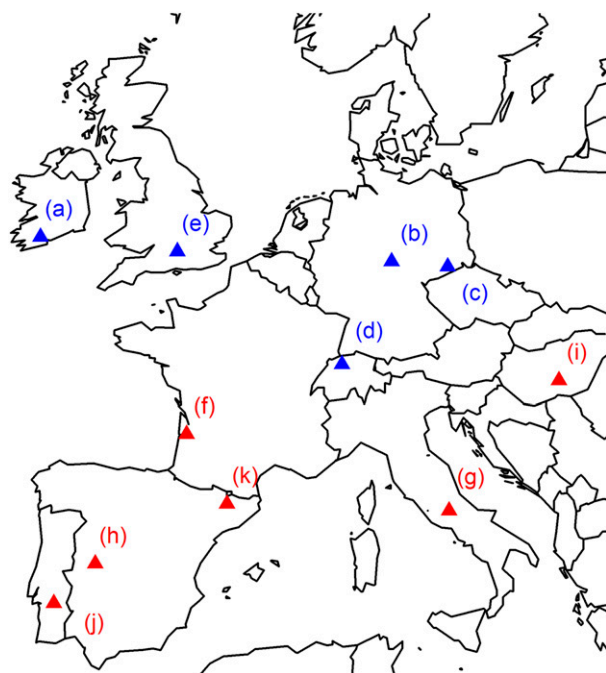
2. Data and methods

a. Dataset

We employed the FluxNet dataset (Baldocchi et al. 2001) for this study using 11 stations situated across Europe. The selection of sites was based both on an initial review of data quality and length across many sites as well as the recommendations of Rebel et al. (2012). Table 1 provides a summary of the site characteristics. To aid the interpretation of the results, we classify the sites as wet or dry based on values of soil moisture. Locations are provided in Fig. 1. At each site, soil moisture measurements from the top 30 cm of soil are provided along with precipitation data as well as the variables required for the calculation of PET via the reference crop Penman–Monteith equation, as described in Zotarelli et al. (2010). These variables include incoming solar radiation, temperature, wind speed, and relative humidity. Among the selected sites, two general land cover types are available: grassland and forest. The data used here are at a daily resolution. We use soil moisture values for the summer months of JJA. For the contributing meteorological variables, we used observations that extend back into previous months in order to calculate integration periods prior to a given soil moisture observation.

b. Conceptual model

We design a conceptual model, based on a framework developed by Bevacqua et al. (2017), in which we describe soil moisture h as an impact of contributing



▲ Wet Sites ▲ Dry Sites

FIG. 1. Locations of FluxNet sites employed for this study.

meteorological variables Y . The contributing meteorological variables include a short-term precipitation variable (Y_{1ps}), a long-term precipitation variable (Y_{2pl}), and a PET variable (Y_{3PET}) that are integrated over periods L_1 , L_2 , and L_3 , respectively. A schematic representation of the variables modeled is given in Fig. 2.

The quantities Y_{1ps} and Y_{2pl} respectively represent the most recent and antecedent precipitation that influence the short- and long-term variability of soil moisture. Their respective integration periods L_1 and L_2 are nonoverlapping. Two precipitation variables are required to better capture the temporal distribution of precipitation that would otherwise be lost using one long-term integration only.

The quantity Y_{3PET} represents PET integrated over the period L_3 . PET is often employed as an estimate of ET in drought indices given the lack of ET data. We calculate PET using the reference crop Penman–Monteith equation as defined in Zotarelli et al. (2010), where it is derived from incoming solar radiation, temperature, wind, and the actual and saturation vapor pressures. The quantity Y_{3PET} includes temperature within its calculation and so can capture heat waves that influence the drying of soil moisture. Depending on the question at hand, the integration length L_3 is varied; more details of this are given in section 2e.

c. Copula

A copula is a multivariate distribution function that describes the dependence structure between random variables independent of their marginal behavior. The selection of structure of dependence, defined by the given copula family, is hence not constrained by the choice of the marginal distribution functions. This feature provides much flexibility in modeling multivariate distributions as it allows for the application of complex marginal distributions (Salvadori et al. 2007). According to Sklar's theorem (Sklar 1959), the joint cumulative distribution function (CDF) F of an n -dimensional random vector $\mathbf{Y} = (Y_1, \dots, Y_n)$ with marginal CDFs F_1, \dots, F_n can be written as

$$F(y_1, \dots, y_n) = C(u_1, \dots, u_n), \quad (1)$$

where C is an n -dimensional copula and $u_i = F_i(y_i)$ are uniformly distributed variables in the domain $[0, 1]$. Provided the marginal distributions F_i are continuous, the multivariate probability density function (PDF) may be decomposed as

$$f(y_1, \dots, y_n) = f_1(y_1) \times \dots \times f_n(y_n) \times c[F_1(y_1), \dots, F_n(y_n)], \quad (2)$$

where c is the copula density and f_i are marginal PDFs.

There exists a large number of bivariate copula families that each provide an explicit formulation for a given structure of dependence. However, the number of copula families applicable to a dimension of three or higher is quite limited (Aas et al. 2009) and in contrast to reality, where heterogeneous dependence structures often exist, each copula will usually assume the same structure of dependence between all marginals (Aas et al. 2009; Acar et al. 2012; Noh et al. 2013; Bevacqua et al. 2017). We therefore employ PCCs that provide higher flexibility than multivariate copulas and more simplicity in terms of the selection of dependence structure (Aas et al. 2009; Noh et al. 2013).

d. Pair copula constructions

PCCs, initially proposed by Joe (1997), allow us to mathematically decompose an n -dimensional copula density into a product of $n(n-1)/2$ bivariate copulas, of which some are conditional. They allow much flexibility in modeling multidimensional distributions (Aas et al. 2009; Bevacqua et al. 2017) and provide a means to easily calculate quantiles of the multivariate conditional distribution of an impact h given values of Y (Noh et al. 2013; Bernard and Czado 2015; Kraus and Czado 2017; Fischer et al. 2017).

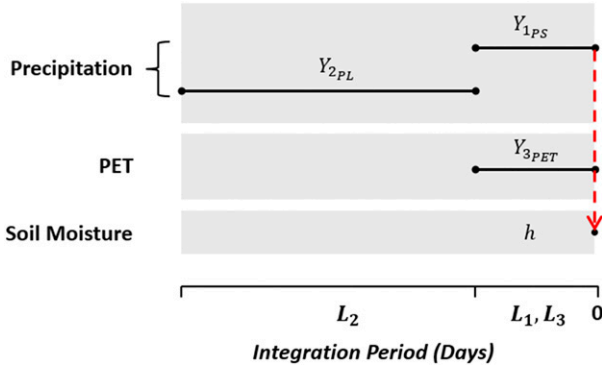


FIG. 2. Schematic of the variables used in this study to construct the soil moisture model.

For a high-dimensional distribution, there exists a significant number of decompositions of a multivariate PDF into a PCC that are each mathematically equivalent to one another (Aas et al. 2009). Two special types of decompositions called vines exist for PCCs, the canonical vine (C-vine) and the D-vine (Kurowicka and Cooke 2005). Throughout this study we employ a D-vine decomposition. For the four-dimensional distribution under study here, there are 12 possible D-vine decompositions. For convenience, we select one decomposition to be applied throughout the study at all sites; the procedure we follow for this selection is outlined in section 2e. The selected D-vine decomposition for the conditional model is given as

$$f_{3,2,1,h}(y_3, y_2, y_1, h) = f_3(y_3) \times f_2(y_2) \times f_1(y_1) \times f_h(h) \\ \times c_{32}(u_3, u_2) \times c_{21}(u_2, u_1) \times c_{1h}(u_1, u_h) \times c_{31|2}(u_{3|2}, u_{1|2}) \\ \times c_{2h|1}(u_{2|1}, u_{h|1}) \times c_{3h|21}(u_{3|21}, u_{h|21}). \quad (3)$$

The differences between each of the 12 possible decompositions are in the ordering of variables within the PCC, which determines the bivariate dependencies that are modeled. As can be seen in Eq. (3), the ordering of variables in the selected decomposition are (Y_3, Y_2, Y_1, h) . To sample h conditioning on the Y , we employ a sampling algorithm provided by the CDVineCopulaConditional R package (Bevacqua 2017), which uses a modified version of the algorithm presented in Aas et al. (2009). This algorithm requires that h is positioned last (or equivalently first) in the order of variables as shown above. This constraint reduces the number of possible decompositions to six. And although each of these possibilities are mathematically equivalent, the explanatory power of h from the resulting conditional

model varies depending on the order of the contributing variables Y within the decomposition, as each Y can have differing levels of influence over h (Kraus and Czado 2017).

1) ESTIMATION OF PCC

The estimation of the PCC given in Eq. (3) is obtained through a sequential approach. First, the unconditional bivariate copulas c_{32} , c_{21} , and c_{1h} are fitted to capture the respective pairwise dependencies of the variables u_3 , u_2 , u_1 , and u_h . Second, the conditional bivariate copulas $c_{31|2}$ and $c_{2h|1}$ are then fitted to the respective conditional probabilities $u_{3|2}$, $u_{1|2}$, $u_{2|1}$, and $u_{h|1}$. These variables are obtained from the conditional distributions given by the partial differentiation of the respective unconditional bivariate copula with respect to the conditioning variable:

$$u_{3|2} = F_{3|2}(u_3|u_2) = \frac{\partial}{\partial u_2} C_{32}(u_3, u_2), \\ u_{1|2} = F_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C_{21}(u_2, u_1), \\ u_{2|1} = F_{2|1}(u_2|u_1) = \frac{\partial}{\partial u_1} C_{21}(u_2, u_1), \quad \text{and} \\ u_{h|1} = F_{h|1}(u_h|u_1) = \frac{\partial}{\partial u_1} C_{1h}(u_1, u_h). \quad (4)$$

In the final step of the estimation procedure, a copula $c_{3h|21}$ is fitted to the conditional probabilities $u_{3|21}$ and $u_{h|21}$. These conditional probabilities are obtained from the conditional distributions given from the partial differentiation of the respective conditional bivariate copula with respect to the conditioning variable:

$$u_{3|21} = F_{3|21}(u_3|u_2, u_1) = \frac{\partial C_{31|2}[F_{3|2}(u_3|u_2), F_{1|2}(u_1|u_2)]}{\partial F_{1|2}(u_1|u_2)} \\ \text{and} \\ u_{h|21} = F_{h|21}(u_h|u_2, u_1) = \frac{\partial C_{2h|1}[F_{2|1}(u_2|u_1), F_{h|1}(u_h|u_1)]}{\partial F_{2|1}(u_2|u_1)}. \quad (5)$$

From the conditional copula $c_{3h|21}$, the conditional CDF $F_{h|321}$ can be obtained through partial differentiation of $C_{3h|21}$ with respect to $F_{3|21}$:

$$F_{h|321}(u_h|u_3, u_2, u_1) \\ = \frac{\partial C_{3h|21}[F_{3|21}(u_3|u_2, u_1), F_{h|21}(u_h|u_2, u_1)]}{\partial F_{3|21}(u_3|u_2, u_1)}. \quad (6)$$

As will be shown, all unconditional and conditional bivariate CDFs described above are required when sampling from the PCC.

2) SAMPLING FROM PCC

Sampling variables u_3 , u_2 , u_1 , and u_h from the four-dimensional D-vine PCC repeatedly results in four uniformly distributed variables that exhibit a dependence structure specified by the given PCC. Algorithms proposed by Aas et al. (2009) provide a convenient means of sampling variables u_3 , u_2 , u_1 , and u_h . Within these algorithms, variables w_3 , w_2 , w_1 , and w_h are first drawn independently from a random uniform distribution on $[0, 1]$. Then, u_3 , u_2 , u_1 , and u_h are determined as

$$\begin{aligned} u_3 &= w_3, \\ u_2 &= F_{2|3}^{-1}(w_2|u_3), \\ u_1 &= F_{1|2}^{-1}[F_{1|23}^{-1}(w_1|u_2, u_3)], \quad \text{and} \\ u_h &= F_{h|1}^{-1}[F_{h|21}^{-1}[F_{h|321}^{-1}(w_h|u_3, u_2, u_1)]]. \end{aligned} \quad (7)$$

Given specified values of Y , the model may be used to sample $h = F_h^{-1}(u_h)$ from a conditional distribution defined by the given Y values. In this case, the variables u_3 , u_2 , u_1 , and u_h are obtained as

$$\begin{aligned} u_3 &= F_3(y_3), \\ u_2 &= F_2(y_2), \\ u_1 &= F_1(y_1), \quad \text{and} \\ u_h &= F_{h|1}^{-1}[F_{h|21}^{-1}[F_{h|321}^{-1}(w_h|u_3, u_2, u_1)]]. \end{aligned} \quad (8)$$

Throughout this study, we use an algorithm proposed by Bevacqua et al. (2017) to sample from Eq. (8) and carry out all simulations using the CDVineCopulaConditional R package (Bevacqua 2017).

When sampling from Eq. (8) given an observed Y , we produce a stochastic time series of h . Repeated simulations conditioning on the observed Y will produce multiple time series with varying statistics and agreement with observed h values (Pham et al. 2016). Throughout this study, given an observed time series of Y , we produce an ensemble consisting of 1000 members of h time series and obtain a probabilistic forecast of h at each time step.

e. Model construction

In this section, we lay out the procedure taken for selecting integration period lengths L_i for the contributing meteorological variables Y_i (Fig. 2). We also provide details of the selection procedure for the D-vine decomposition of the PCC and the selection of copula families within the PCC.

1) METEOROLOGICAL PREDICTOR SELECTION

We describe soil moisture h as a function of two precipitation variables, $Y_{1\text{ps}}$ and $Y_{2\text{pl}}$, integrated over periods L_1 and L_2 , and a PET variable, $Y_{3\text{pet}}$ integrated over the period L_3 . By developing a statistical model with these variables and soil moisture, we look to answer the following three questions:

- 1) What are the individual contributions of the meteorological variables Y_i to the estimation of soil moisture h on time scales related to meteorological drought and heat waves?
- 2) What relevance does the dependence between antecedent precipitation ($Y_{2\text{pl}}$) and recent PET ($Y_{3\text{pet}}$) have for the estimation of low soil moisture values?
- 3) What relevance does PET have for the estimation of soil moisture over varying integration lengths L_3 ?

To answer these questions, we propose two sets of Y variables, S1 and S2. Questions 1 and 2 are then approached using variable set S1 while Question 3 is approached using variable set S2. The difference between S1 and S2 is the integration L_3 chosen at each site. A short integration period is considered for PET in S1, while a long integration period is considered for PET in S2. For each value of L_i used, the contributing meteorological variable Y_i may be defined as

$$\begin{aligned} Y_{1\text{ps}}(t) &= \sum_{t-L_1+1}^t p(t), \\ Y_{2\text{pl}}(t) &= \sum_{t-(L_1+L_2)+1}^{t-L_1} p(t), \quad \text{and} \\ Y_{3\text{pet}}(t) &= \sum_{t-L_3+1}^t \text{pet}(t), \end{aligned} \quad (9)$$

where $p(t)$ and $\text{pet}(t)$ are daily precipitation and PET, respectively.

We address the first two questions with variable set S1. The selected L_i for S1 must result in Y variables that provide satisfactory estimates of soil moisture h , hold physically meaningful dependencies, and capture time scales relevant for both meteorological drought and heat waves. Physically meaningful dependencies are obtained by constraining L_i such that $L_1 = L_3$ and through ensuring that there is no overlap between L_2 and the short-term integrations.

Based on the analysis described below, we find a difference between grassland sites and forest sites. Forest sites require a longer integration L_1 . This is possibly explained by the deeper root systems at forest sites, which filter the influence of short-term variability in rainfall on the integrated soil column. We therefore

choose two sets of L : LG and LF, for grassland and forest sites, respectively. At all grassland (forest) sites, the same LG (LF) are used.

We choose integrations of $LG_1 = LG_3 = 7$ and $LG_2 = 63$ for grassland sites. For forest sites, we choose integrations of $LF_1 = LF_3 = 30$ and $LF_2 = 60$. We thus use information of precipitation over the previous 70 and 90 days for each daily soil moisture observation at grassland and forest sites, respectively.

To select LG_i (LF_i) in S1, we first calculate the Spearman correlation between $Y_i(t)$ and $h(t)$ for multiple integrations within a window of 120 days prior to day t . We then choose the integration length that maximizes the Spearman correlation for each Y_i . Integration periods are then constrained such that $LG_1 = LG_3$ ($LF_1 = LF_3$). This ensures physically meaningful dependencies and avoids arbitrary dependencies that would otherwise arise between differing LG_1 (LF_1) and LG_3 (LF_3).

The sensitivity of the conditional model's performance, in representing h conditioning on Y , to changes in LF (LG) is tested by varying the short-term LG (LF) by ± 4 days while the long-term integration LG (LF) is varied by ± 10 days. Changes in performance are found to be minimal (not shown). Assuming the same LG (LF) at all grassland (forest sites) and constraining the integration periods is therefore expected to have little weight in the outcome of this analysis.

We acknowledge in S1 that the influence of most recent daily temperature extremes on soil moisture is potentially filtered out at forest sites by setting $LF_3 = 30$. This is addressed in variable set S2 where we assess the relevance of the selection of L_3 to the estimation of h (question 3). In S2, two models are constructed using a short- and long-term integration of L_3 . The same LG_1 , LF_1 , LG_2 , and LF_2 as S1 are used while LG_3 and LF_3 are set to 7 and 70 days and 7 and 90 days, respectively.

As the variables are all calculated on a daily resolution, from day t to day $t + 1$, there will be an overlap of $LG_i - 1$ or $LF_i - 1$ mutual days used in the integration periods associated with two consecutive days. We thus violate the assumption that data are independent and identically distributed, which the statistical methods used here are based upon. It should therefore be noted that the performance of the model as well as any estimated dependence between variables may be overestimated.

2) STATISTICAL INFERENCE OF THE MULTIVARIATE PDF

The parameters of each bivariate copula in Eq. (3) are estimated based on the marginal variables u_i drawn from the marginal CDFs F_i . We use a kernel density

estimate for all marginal distributions. All marginal densities are estimated using the *ks* R package (Duong 2017), which employs the bandwidth selector of Wand and Jones (1994).

The estimation of copula parameters requires that no equal ranks are present in u_i . We follow the approach used in Pham et al. (2016) to remove ties from the data. In this approach, a small random noise is drawn from a uniform distribution on $[-0.001, 0.001]$ and added to Y_{1ps} and Y_{2pl} values greater than zero. For values equal to zero, we add a random noise drawn from the uniform distribution on $[0, 0.001]$.

The use of kernel density estimates provides a convenient way of estimating the marginal distribution of h . Soil moisture has natural upper and lower bounds, according to its wilting and saturation points, respectively, and can also exhibit a bimodal distribution (Porporato and D'Odorico 2004; D'Andrea et al. 2006).

The selection of the D-vine decomposition in Eq. (3) is based on an initial test in which we assess the performance of each of the six possible decompositions in their ability to represent h when conditioning on the observed Y . At all sites we fit a PCC for each of the six decompositions and use the Akaike information criterion (AIC) when selecting the type of copulas to be used. The selection of copula families and the estimation of their parameters is carried out at each site separately. Each copula is chosen from a range of copulas provided by the VineCopula R package (Schepsmeier et al. 2017). To assess each of the six possible decompositions, a probabilistic forecast of h consisting of 1000 members is produced at all sites. These are compared with observed soil moisture using the root-mean-square error. We then select the decomposition that generally shows the highest explanatory power of h at all sites.

After selecting the decomposition to apply, the goodness of fit (GoF) of the selected copulas is tested. Copulas initially selected according to the AIC did not always provide a satisfactory fit. For this reason we use two criteria in the selection of a copula for each pair in the PCC. This procedure is carried out sequentially as outlined in section 2d(1), where unconditional copulas are first selected followed by the conditional copulas. For each pair, we first select the top three copulas according to the AIC and second test the GoF of each using K plots (Genest and Favre 2007; Bevacqua et al. 2017). We then select the highest ranked copula according to the AIC that shows satisfactory compliance in the K plots.

A K plot is a plot of the Kendall function $K(w) = P[C_{i,j}(U_i, U_j) \leq w]$ obtained from the fitted copula against $K(w)$ computed with the empirical copula obtained

using the observed data. Similarly to a Q–Q plot for univariate distributions, the K plot indicates good quality of fit when points follow the diagonal. These plots provide uncertainties around the empirical copula as well as a qualitative idea of the quality of fit of each copula (Bevacqua et al. 2017). Most selected copulas show good agreement according to the K plots (not shown) where parametric $K(w)$ values generally follow the mean of the empirical values and mostly remain within the uncertainty intervals calculated from 1000 simulations. Some small problems are found with the copulas at sites e and f, which may limit the strength of conclusions drawn from these sites.

f. Model evaluation metrics

The model simulations are evaluated overall and in their ability to represent low values of soil moisture h . Using the Brier score (BS), we evaluate the accuracy of probabilistic predictions of low h values defined as those below the 15th percentile of observed soil moisture. The closer BS is to zero, the better the predictions. The BS is defined as:

$$\text{BS} = \frac{1}{N} \sum_{t=1}^N (p_t - o_t)^2, \quad (10)$$

where p_t is the probability of getting a simulated value of h below the observed 15th percentile from the model at time t , while o_t is 1 if observed soil moisture $h^{\text{obs}}(t)$ is below the 15th percentile and 0 otherwise. Along with BS we calculate the associated Brier skill score (BSS) that evaluates the model relative to a reference model BS_{ref} :

$$\text{BSS} = 1 - \frac{\text{BS}}{\text{BS}_{\text{ref}}}, \quad (11)$$

We consider the climatology as the reference model in which the probability of a value occurring below the 15th percentile is always 0.15.

The model is also evaluated in its ability to capture the persistence of drought conditions by comparing the autocorrelation function (ACF) and using an empirical order 1 persistence probability (PP). Both are derived from the observed values and the mean of the simulated values. We choose an order 1 persistence after assessing partial autocorrelation function (PACF) at each site, which only showed significant correlations for order 1. The quantity PP is defined as

$$\text{PP} = \Pr[h_{t+1} < F_h^{-1}(0.15) | h_t < F_h^{-1}(0.15)]. \quad (12)$$

The PP may be interpreted as the probability that h_{t+1} will be below the 15th percentile given that h_t is below the 15th percentile.

3. Results

The set of variables S1, described in section 2e(1), are employed to evaluate the contributions of the individual Y variables and that of their dependence structure to soil moisture. To achieve this we perform a number of sensitivity simulations and compare them with a control simulation (CTRL). All simulations carried out are done through a K-fold cross validation to avoid overfitting. Parameter K here is the number of summers in a time series at a given site. In each simulation, we thus remove one summer at a time when fitting the copula parameters but use the same marginal PDFs for each period. In this way we only cross-validate the PCC rather than the entire multivariate statistical model. For each simulation, we then produce a probabilistic forecast of h consisting of 1000 members through conditioning on specified values of Y .

a. Model performance

The CTRL simulation is performed through sampling h conditioned on observed values of Y [Eq. (8)]. The performance of CTRL may be qualitatively gauged from Fig. 3. Plots shown in Figs. 3a–e are results from wet sites while those from Figs. 3f–k are results from dry sites. The mean value of h from CTRL at each time step can be seen to follow the temporal evolution of observed soil moisture (h^{obs}) quite well, while h^{obs} is generally found within the 95% confidence interval of CTRL. Also shown within each panel in Fig. 3 are the order 1 persistence probabilities of low h for observed (PP_{obs}) and mean simulated h (PP_{sim}). The quantities PP_{sim} and PP_{obs} are found to be very similar at all wet sites and most dry sites, although PP_{sim} is generally less than PP_{obs} at dry sites. A comparison of the observed ACF, estimated up to order 10, with the ACF derived from the mean of the simulation also showed close correspondence at each site (not shown). Such results indicate good agreement between the observed h and simulated mean h in terms of temporal evolution and the persistence of low values.

To provide information of the performance of the model in terms of the probabilistic forecast, we calculate BS and BSS for CTRL at each site (Table 2). In general we see good BS and positive BSS that range from 0.06 to 0.12 and 0.04 to 0.51, respectively, with medians of 0.09 and 0.25. These BSS indicate that the model is better than the climatology at predicting low soil moisture values. Low BSS values are seen at site c, where we also see poor correspondence between h^{obs} and the mean of CTRL. Optimizing the performance of the model at this site through changing integration periods does not bring a noticeable improvement,

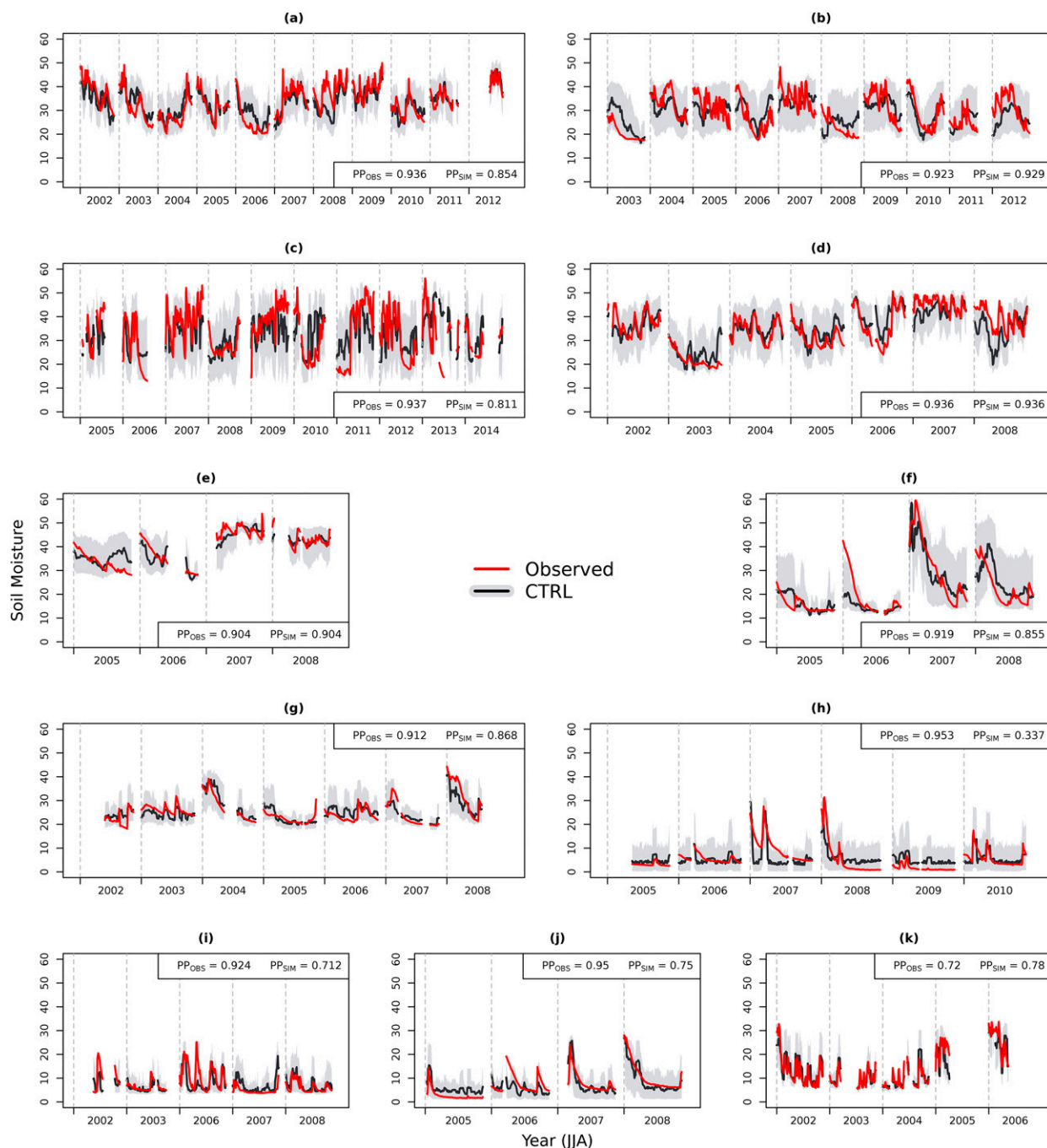


FIG. 3. Observed time series (red) alongside the cross-validation time series of the CTRL mean (black) and the 95% prediction interval (gray), obtained from 1000 simulations, at (a)–(e) wet sites and (f)–(k) dry sites. Also provided within each panel are the order 1 persistence probabilities calculated from the observed (PP_{obs}) and CTRL mean (PP_{sim}) time series.

indicating that the proposed model and variables included do not predict soil moisture correctly at all sites. However, with satisfactory results generally obtained at most sites, we employ the model for use in sensitivity analysis in a number of tests presented below.

b. Assessment of contributing variables to soil moisture

We test the contribution of Y_{1ps} (short-term precipitation), Y_{2pl} (long-term precipitation), and Y_{3pet} (PET) to the estimation of h in three sensitivity simulations

TABLE 2. BS, BSS, and mean bias for CTRL, SENS- Y_{1PS} , SENS- Y_{2PL} , and SENS- Y_{3PET} simulations calculated for soil moisture values below the observed 15th percentile. Bias values for SENS- Y_{1PS} , SENS- Y_{2PL} , and SENS- Y_{3PET} are given as percentage change relative to CTRL.

Site	Score	CTRL	SENS- Y_{1PS}	SENS- Y_{2PL}	SENS- Y_{3PET}
a	BS	0.09	0.10	0.12	0.10
	BSS	0.25	0.21	0.01	0.18
	Bias	3.89	+5%	+113%	+45%
b	BS	0.11	0.13	0.10	0.11
	BSS	0.15	-0.01	0.17	0.15
	Bias	4.33	+107%	-1%	+33%
c	BS	0.12	0.14	0.11	0.12
	BSS	0.04	-0.1	0.13	0.06
	Bias	10.66	+51%	-24%	+2%
d	BS	0.06	0.06	0.09	0.08
	BSS	0.51	0.49	0.26	0.39
	Bias	3.52	+45%	+149%	+83%
e	BS	0.09	0.10	0.09	0.12
	BSS	0.28	0.25	0.27	0.03
	Bias	3.62	-15%	+78%	+81%
f	BS	0.08	0.07	0.06	0.13
	BSS	0.36	0.43	0.53	-0.01
	Bias	0.37	-215%	+207%	+720%
g	BS	0.09	0.09	0.11	0.09
	BSS	0.31	0.24	0.15	0.30
	Bias	1.48	+43%	+86%	+6%
h	BS	0.12	0.13	0.12	0.13
	BSS	0.04	0.00	0.05	0.003
	Bias	3.28	+9%	-3%	+4%
i	BS	0.12	0.13	0.12	0.13
	BSS	0.07	0.002	0.04	-0.06
	Bias	1.24	-9%	+5%	+52%
j	BS	0.12	0.12	0.12	0.12
	BSS	0.09	0.10	0.05	0.06
	Bias	2.8	-8%	+10%	+16%
k	BS	0.08	0.13	0.12	0.08
	BSS	0.36	0.01	0.09	0.37
	Bias	1.05	+205%	+146%	+10%

SENS- Y_{1PS} , SENS- Y_{2PL} , and SENS- Y_{3PET} , respectively. For each sensitivity simulation, h is sampled conditioning on the median value of the respective variable to be tested and the observed values of the other two Y variables. To assess the contributions of all variables, we compare the mean of each simulation with the CTRL mean. We also compare the probabilistic forecasts from SENS- Y_i with CTRL using the BS, BSS, and the mean ensemble bias computed for values of h below the observed 15th percentile (Table 2).

At wet sites, precipitation is generally seen to have the most influence on low soil moisture values, while PET can act to amplify the low soil moisture anomaly during

drought periods. Comparing the means of the three sensitivity simulations with the mean of CTRL (Fig. 4), larger overestimations of low h values with respect to CTRL are generally seen in either of the simulations assessing the influence of a precipitation variable, SENS- Y_{1PS} or SENS- Y_{2PL} , than is, seen in SENS- Y_{3PET} . Underlining this are larger changes in positive bias of low soil moisture values seen from SENS- Y_{1PS} or SENS- Y_{2PL} than from SENS- Y_{3PET} (Table 2). A comparison of BSS for each simulation in Table 2 also shows a larger reduction in skill of forecasting values below 15th percentile in either SENS- Y_{1PS} or SENS- Y_{2PL} than in SENS- Y_{3PET} . Focusing on drought events at wet sites a, b, and d in 2003 and 2006, years in which heat waves have also occurred (Ciais et al. 2005; Rebetez et al. 2009), we see from the mean of the simulations (Fig. 5) that removing the influence of precipitation can lead to the misspecification of a drought event with the green line largely above the black line (CTRL). On the other hand, removing the influence of PET can result in the underestimation of the severity of the event, with the blue line only just higher than the black during a drought event.

At dry sites, we see that precipitation again holds the main influence over soil moisture while PET generally offers little added benefit to the estimation of soil moisture. The main differences of CTRL with SENS- Y_{1PS} and SENS- Y_{2PL} are found for high values of soil moisture (Fig. 4). Low values in these sensitivity simulations are generally equivalent with CTRL, as the medians of Y_{1PS} and Y_{2PL} are associated with relatively low values due to the positively skewed nature of the variables' distributions. Little or no difference is seen between SENS- Y_{3PET} and CTRL simulations for low values of soil moisture. Large percentage changes in bias for low soil moisture values are seen at sites f and i, though the actual changes in soil moisture are relatively low (Table 2). This would be expected at dry sites during summer where soil moisture normally reaches low levels such that ET is moisture-limited and will diverge from PET. Extremes of PET driven by extreme temperatures would then have little added effect to the severity of soil moisture drought in dry locations.

c. Assessing the relevance of Y dependence structure

The contribution of the dependence between Y_{2PL} and Y_{3PET} to the estimation of low h values is assessed using the sensitivity simulation IND- Y_{2PL} , which is used to highlight where interactions between drought and heat wave conditions, arising through land-atmosphere interactions, act to amplify drought conditions. To illustrate the dependence between Y_{2PL} and Y_{3PET} , we calculate Spearman's ρ and a measure of tail dependence λ_q , calculated as

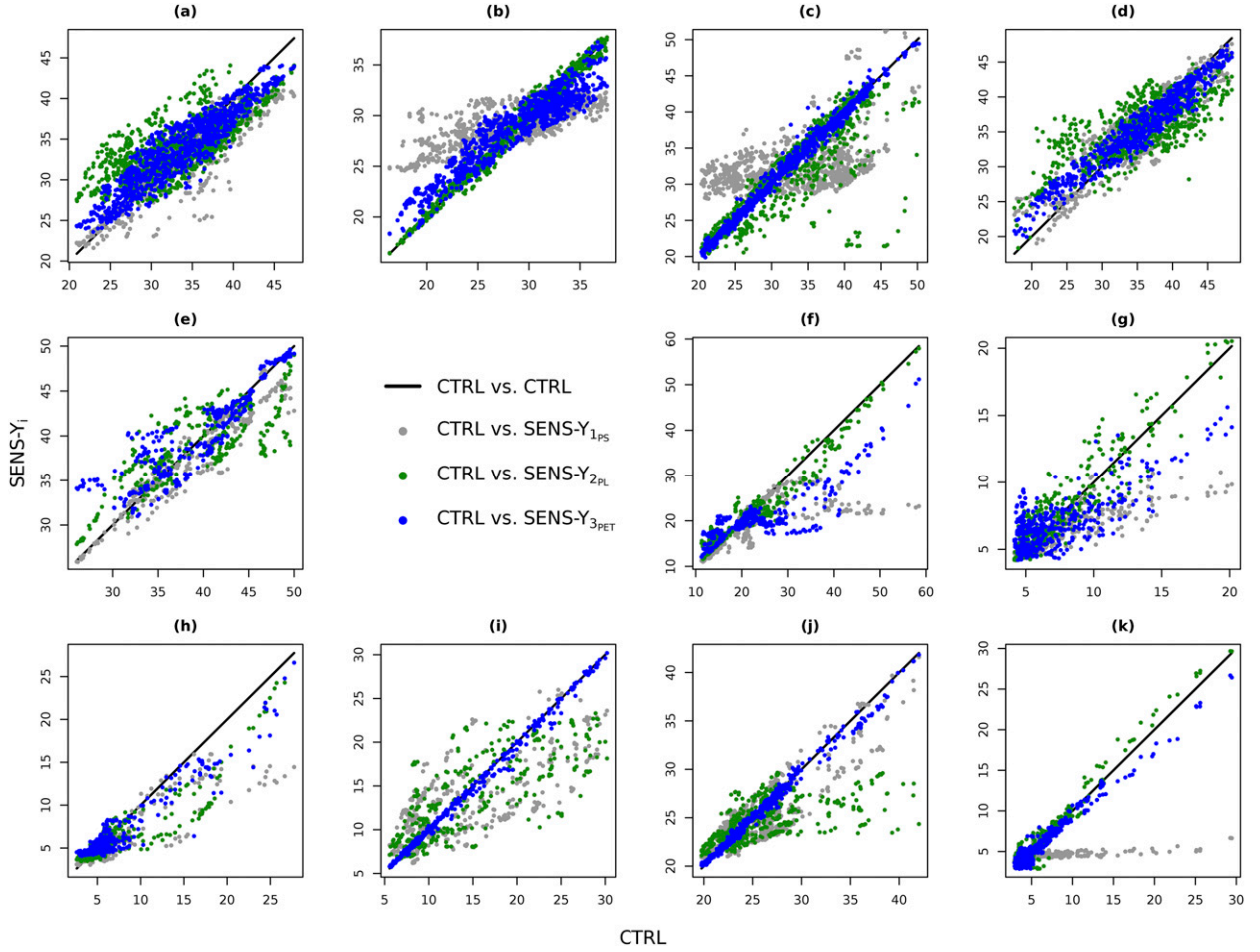


FIG. 4. Comparison of the mean of the cross-validation simulations of CTRL with SENS- Y_{1ps} (gray dots), SENS- Y_{2pl} (green dots), and SENS- Y_{3pet} (blue dots) at (a)–(e) wet sites and (f)–(k) dry sites. Values are ordered according to CTRL from low to high such that the closer the correspondence of points to the diagonal, the smaller the change in the estimation of soil moisture in the given sensitivity simulation.

$$\lambda_q = \Pr \left[Y_{3pet} > F_3^{-1}(q) \mid Y_{2pl} < F_2^{-1}(1-q) \right], \quad (13)$$

where $q = 0.9$ in this case. The quantity λ_{90} can be interpreted as the fraction of days when Y_{3pet} was greater than its observed 90th percentile when Y_{2pl} was less than its 10th percentile. For two independent variables, the expected value of λ_q is $1 - q$. Values of ρ and λ_q for each site are given in Fig. 6. At many sites we observe a negative dependence between Y_{2pl} and Y_{3pet} , as measured by ρ , and an increased probability of extreme PET (Y_{3pet}) when antecedent precipitation (Y_{2pl}) had been extremely low.

To test the relevance of such dependence in IND- Y_{2pl} , we break the dependence between Y_{2pl} and the short-term variables Y_{1ps} and Y_{3pet} . This is achieved by shuffling Y_{2pl} such that it is randomly associated with them. A probabilistic forecast of h , consisting of

1000 members, is then produced, sampling from the multivariate distribution where we condition on the observed values of Y_{1ps} and Y_{3pet} and the shuffled Y_{2pl} . To account for sampling variability of the shuffling process, we produce 1000 IND- Y_{2pl} probabilistic forecasts.

We obtain a kernel density estimate of the PDF produced from each of the 1000 IND- Y_{2pl} simulations. The mean density and the 95% confidence interval of IND- Y_{2pl} PDFs are calculated and presented alongside the PDFs of CTRL and h^{obs} (Fig. 6). The statistical significance of the difference between the CDFs of CTRL and IND- Y_{2pl} is assessed at the 5th, 10th, and 15th percentiles of observed soil moisture. CTRL is considered significantly different for a given percentile if the associated soil moisture value of CTRL is less than the lower bound of 95% confidence interval of that percentile from IND- Y_{2pl} . This would signify that the probability

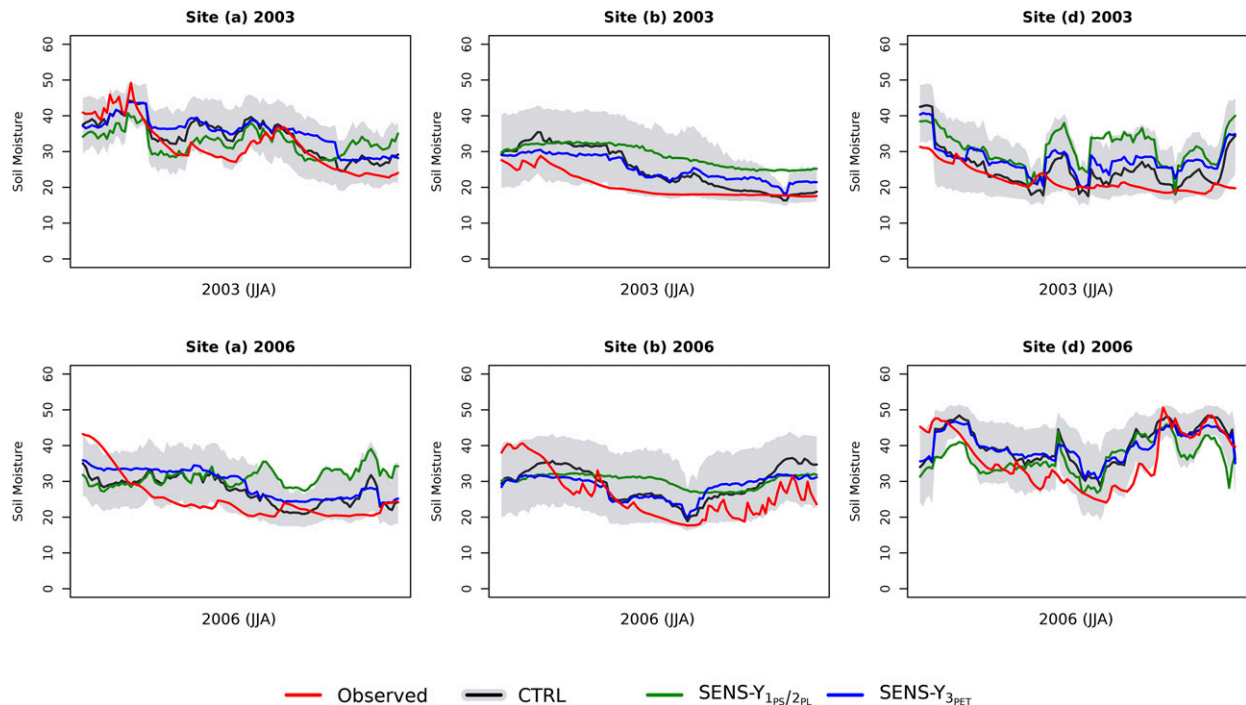


FIG. 5. Mean cross-validated time series of simulations assessing the contributions of precipitation and PET to the estimation of soil moisture and CTRL (black) for the summers (JJA) of (top) 2003 and (bottom) 2006 at wet sites a, b, and, d. Time series of mean simulated values are presented for SENS- Y_{2PL} (green) and SENS- Y_{3PET} (blue) at wet sites a and d while time series of SENS- Y_{1PS} (green) and SENS- Y_{3PET} (blue) are presented for site b.

of values below that percentile are underestimated when the dependence between Y_{2PL} and Y_{3PET} is broken.

Statistically significant differences are found between all three percentiles at site d, where we also see a noticeable difference between PDFs (Fig. 6d). A negative dependence as well as a significant dependence in the tails is also observed here. Site d lies in a transitional region where land–atmosphere interactions can lead to the mutual reinforcement of drought and heat wave events (Seneviratne et al. 2010). This result highlights the importance of the interplay between drought and heat wave conditions, driven by land–atmosphere interactions, to the reinforcement of drought conditions in such locations.

Statistically significant differences between the percentiles tested are also found at wet sites a, b, and e and dry sites g and j, though relatively little difference is observed between CTRL and IND- Y_{2PL} PDFs at these sites for values below the tested percentiles (Fig. 6). We observe negative dependencies (ρ) and tail dependencies (λ_q) at these sites, which highlights that the concurrence of such conditions may be important for the estimation of low values of soil moisture. These dependencies are also observed at other dry sites, but no significant differences between assessed percentiles are found. Such dependencies at these sites are perhaps

of little relevance for soil moisture during summer, as extremes of PET may be energy limited in wet climates while soil in dry climates may have little available moisture for ET. In dry conditions then, extremes of PET in combination with extremely low antecedent precipitation will have little effect on moisture levels in soil.

d. Relevance of PET over short and long integration periods

The variable set S2, as described in section 2e(1), is used to demonstrate the relevance of PET, integrated over various durations LG_3 and LF_3 , to the estimation of soil moisture h . We fit two models at wet sites a, b, and d, where we see contributions of PET to the estimation of soil moisture drought in variable set S1 (Fig. 5). The integration periods used for precipitation variables Y_{1PS} and Y_{2PL} in S1 remain the same. For the simulation PET-INT_S, we set $LG_3 = LF_3 = 7$, and for the simulation PET-INT_L, we set $LG_3 = 70$ and $LF_3 = 90$.

Based on the mean of the simulations (Fig. 7), better representation of drought onset can be seen at sites a and d in the PET-INT_S simulations where the black line generally follows red (observed) at the beginning of an event when initial drying is taking place. On the other hand, drought persistence is generally captured better

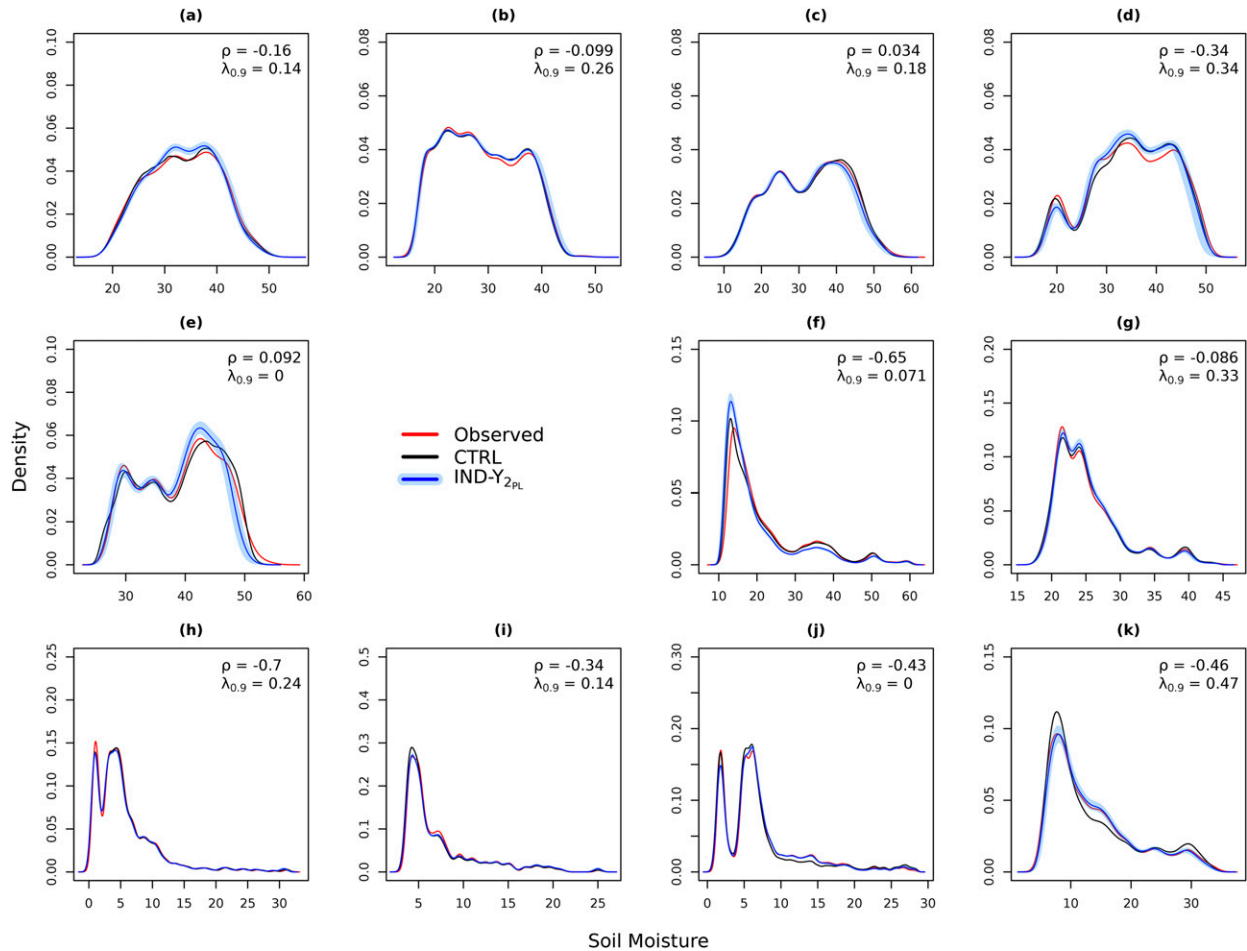


FIG. 6. Kernel density estimates of observed soil moisture (red) and soil moisture simulated via cross-validation from probabilistic forecasts CTRL (black) and $\text{IND-}Y_{2\text{PL}}$ (blue) simulations. The blue line and shading respectively represent the mean density and 95% confidence interval obtained from the 1000 $\text{IND-}Y_{2\text{PL}}$ simulations.

by the PET-INT_L simulation where the blue line remains low with the red line in comparison to the black. Better BSSs are found for simulations using a long-term integration of PET at sites a and b. Increases of BSS, from PET-INT_S to PET-INT_L , from 0.24 to 0.36 and from 0.18 to 0.25 are found at each site, respectively, while little difference is seen between simulations at site d with BSS equal to 0.51 and 0.52.

Although these results are somewhat qualitative, they highlight that both short- and long-term integrations of PET are important for the estimation of drought events in this framework. Longer integrations are generally better in capturing the persistence of drought conditions as they can account for the memory soil moisture holds of drying during the event. Short-term integrations, however, are better in capturing drought onset as they are able to account for short intense periods of drying that can accelerate the propagation of meteorological drought to soil moisture drought. With drought events expected to set in

quicker in a warming climate (Trenberth et al. 2014), it will be important to detect such changes in the intensity of drying over short periods in spring and summer that are filtered out in longer integrations of PET. This may be of particular relevance in Europe, where early onset of drought conditions can have large implications for extreme temperatures in summer (Vautard et al. 2007).

4. Summary and conclusions

Compound events are multivariate extreme events in which the contributing variables need not be extreme themselves, but their joint dependent concurrence produces an extreme impact (Leonard et al. 2014; Bevacqua et al. 2017). We have analyzed soil moisture drought over Europe as a compound event of variables employed in common drought indices, namely, precipitation and PET, and assessed the individual roles of these variables and that of their dependence structure to the estimation of soil

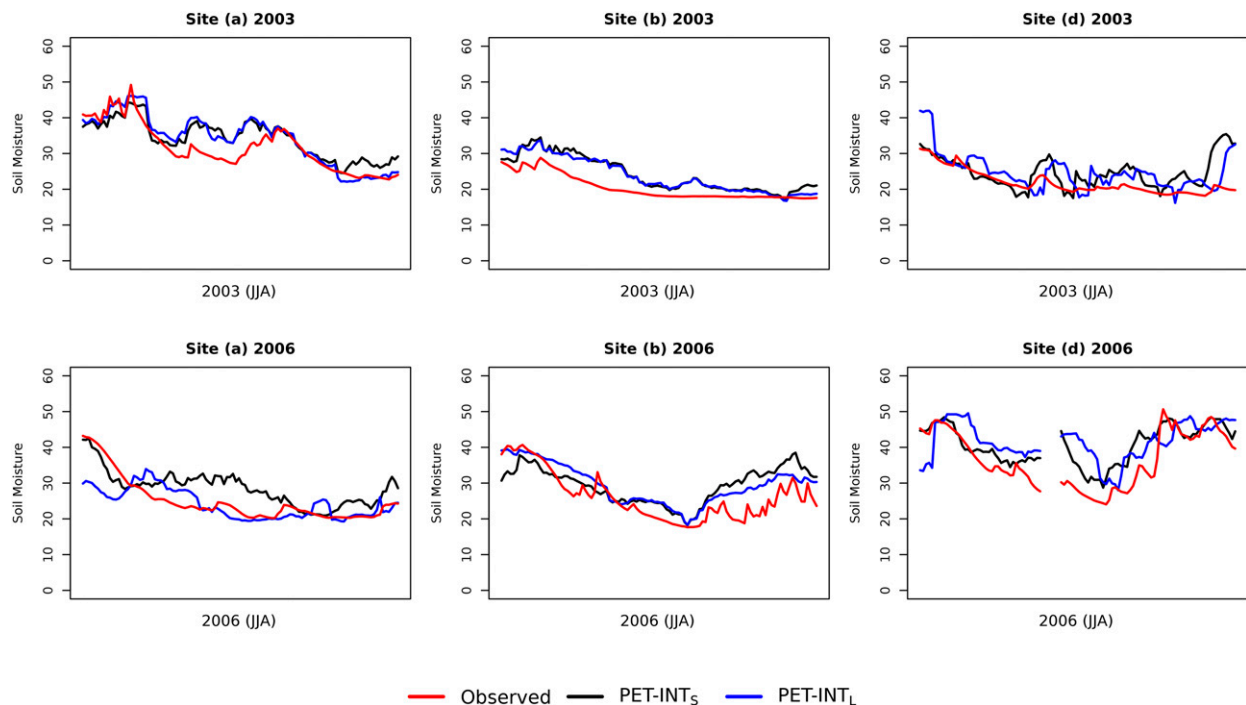


FIG. 7. Mean cross-validation time series of simulations from models PET-INT_S, in which PET is considered over a short integration period (black), and PET-INT_L, in which PET is considered over a long integration period (blue), along with the observed time series (red) for the (top) 2003 and (bottom) 2006 drought events at wet sites a, b, and d.

moisture. The overall aim was to explore the compound nature of soil moisture drought and the differences that exist between wet and dry climates.

To achieve our aim, we developed a statistical model based on pair copula constructions. Within the model we considered precipitation and PET over time scales related to meteorological drought and heat waves, respectively. These time scales were considered to assess the influence of heat wave conditions on soil moisture, as well as dependencies driven by land–atmosphere interactions that can cause a mutual reinforcement between drought and heat wave events in Europe. We applied the model to data from 11 FluxNet sites situated in wet, transitional, and dry climates in Europe and generally found satisfactory performance of the model. We thus employed it in a number of sensitivity experiments to assess the relevance of contributing variables and their dependence structure to the estimation of soil moisture drought.

Results obtained from sensitivity experiments were in line with previous studies. Precipitation was found to hold the main control over soil moisture drought. PET was required only when it departs from normal conditions (Vicente-Serrano et al. 2010) to partly explain the severity of drought conditions in wet climates (Seneviratne et al. 2012b; Teuling et al. 2013), while little or no contribution was found in dry climates (Luo et al. 2017)

during summer. The concurrence of extremely low antecedent precipitation with extremely high PET was found to be most relevant at a site situated in a transitional climate region between wet and dry climates where land–atmosphere interactions are most relevant for the development of soil moisture drought (Seneviratne et al. 2006, 2012a). The concurrence of these conditions was also seen at many dry sites, though they were found to have little relevance for soil moisture. This lack of relevance at dry sites is presumably related to the limited availability of moisture in soil for actual ET to occur, such that PET and extremes of PET could have little influence to a low soil moisture anomaly.

The aforementioned contribution of PET is based on a short-term integration period that was used to capture the influence of heat waves on soil moisture. At wet sites, this short integration period is found to be effective in describing the onset of drought events as it can capture initial drying that occurs on a daily basis. It can, however, be ineffective in capturing the persistence of drought conditions, which longer integrations can better account for, as it neglects the memory soil moisture may hold of PET and the intense drying that may have occurred throughout a drought event. The differences found between short and long integrations of PET may become relevant in the analysis of changes in the

onset of drought events using drought indices. A warmer climate may cause droughts to set in quicker (Trenberth et al. 2014) and lead to flash droughts (Mo and Lettenmaier 2016). Such drying may be hidden through the use of longer integration periods of PET in an index such as the SPEI or through a recursive model used for the PDSI that retains memory of PDSI values from previous time steps.

Advantages of using drought indices include the simplicity they offer and the widespread availability of meteorological datasets compared to those of soil moisture. Although they are not specifically designed to represent soil moisture (Seneviratne et al. 2010), indices such as the SPEI, PDSI, and RDI provide a convenient means of combining precipitation and PET into a kind of impact function that may be implicitly linked to soil moisture.

However, soil moisture drought is not a simple phenomenon to characterize with drought indices due to differing contributions and relevant integration periods of meteorological variables in wet and dry climates. The use of a climatic water balance (precipitation – PET) in the SPEI and PDSI assumes oversimplified relationships between precipitation, PET, and soil moisture (Seneviratne 2012) and implies that the statistical relevance of precipitation and PET to the estimation of soil moisture are the same over a given integration period. With such simplifications comes a loss of information, such as short intense periods of drying that may be filtered out through the inclusion of redundant information when using a long integration period for PET.

Through the inclusion of PET, these indices are expected to provide a better picture of changes in drought conditions in a warming climate than indices that use precipitation alone such as the standardized precipitation index (SPI). Ubiquitously applying indices that incorporate PET across different climates can provide a general overview of the response of drought conditions to global warming. It is, however, important to note that severe drought, as depicted by these indices, will have a different meaning for soil moisture drought in wet and dry climates. ET is limited by moisture availability and so will diverge from PET in dry conditions, leading to an overestimation of the actual drying taking place with respect to soil. In contrast, land surface models account for this moisture limitation by capturing the physical relationship between PET and soil moisture; they can therefore provide a more reliable estimate. Their use within coupled climate models to study changes in soil moisture drought is particularly advocated for by Berg et al. (2017), who also demonstrate the added complexity of diverging changes to soil moisture at different soil depths that cannot be disentangled using drought indices.

Despite discrepancies between PET and ET in dry conditions, extremes of PET will still be indicative of the

drying potential of the atmosphere, provided it is calculated using a reliable physically based method such as the Penman–Monteith equation. Such atmospheric drying potential may possibly have adverse effects on crop yields and contribute to other environmental hazards such as wildfires that are mediated by the availability of moisture in vegetation (Gudmundsson et al. 2014; Ruffault et al. 2016).

Much information of soil moisture and other drought impacts may be deduced from drought indices and their response to a warming climate. To do so requires careful interpretation and detailed knowledge of the involved variables' influence on soil moisture in a given climate. It is therefore important that drought indices incorporating PET are interpreted within the context of the climate in which they are applied, while also keeping in mind the applications they are designed for.

In our impact focused approach, we have made use of the little soil moisture data that are available across different locations and climate types in Europe to demonstrate the compound nature of soil moisture drought during summer. These results provide further insight into the relationship between soil moisture and drought indices that incorporate PET. It is hoped that this insight will aid with the interpretation of drought indices in a given climate and season so that as much information as possible may be gained from their application.

Acknowledgments. Colin Manning received funding from the Volkswagen Foundation's CE:LLO project (Az.: 88469), which also supported project meetings. The authors thank Prof. Arnaldo Frigessi and Prof. Ingrid Hobaek Haff for fruitful discussions and the hosting of a project meeting at the Norwegian Computing Center. The authors would also like to sincerely thank the anonymous reviewers for their constructive comments that have enhanced the quality of the manuscript.

REFERENCES

- Aas, K., C. Czado, A. Frigessi, and H. Bakken, 2009: Pair-copula constructions of multiple dependence. *Insur. Math. Econ.*, **44**, 182–198, <https://doi.org/10.1016/j.insmatheco.2007.02.001>.
- Acar, E. F., C. Genest, and J. Nelehov, 2012: Beyond simplified pair-copula constructions. *J. Multivariate Anal.*, **110**, 74–90, <https://doi.org/10.1016/j.jmva.2012.02.001>.
- AghaKouchak, A., L. Cheng, O. Mazdinyasni, and A. Farahmand, 2014: Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophys. Res. Lett.*, **41**, 8847–8852, <https://doi.org/10.1002/2014GL062308>.
- Baldocchi, D., and Coauthors, 2001: FluxNet: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bull. Amer. Meteor. Soc.*, **82**, 2415–2434, [https://doi.org/10.1175/1520-0477\(2001\)082<2415:FANTTS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2).

- Berg, A., J. Sheffield, and P. C. Milly, 2017: Divergent surface and total soil moisture projections under global warming. *Geophys. Res. Lett.*, **44**, 236–244, <https://doi.org/10.1002/2016GL071921>.
- Bernard, C., and C. Czado, 2015: Conditional quantiles and tail dependence. *J. Multivar. Anal.*, **138**, 104–126, <https://doi.org/10.1016/j.jmva.2015.01.011>.
- Bevacqua, E., 2017: CDVineCopulaConditional: Sampling from conditional C- and D-vine copulas. R package, version 0.1.0, <https://CRAN.R-project.org/package=CDVineCopulaConditional>.
- , D. Maraun, I. Hobk Haff, M. Widmann, and M. Vrac, 2017: Multivariate statistical modelling of compound events via pair-copula constructions: Analysis of floods in Ravenna (Italy). *Hydrol. Earth Syst. Sci.*, **21**, 2701–2723, <https://doi.org/10.5194/hess-21-2701-2017>.
- Ciais, P., and Coauthors, 2005: Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, **437**, 529–533, <https://doi.org/10.1038/nature03972>.
- Dai, A., 2011: Drought under global warming: A review. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 45–65, <https://doi.org/10.1002/wcc.81>.
- , 2013: Increasing drought under global warming in observations and models. *Nat. Climate Change*, **3**, 52–58, <https://doi.org/10.1038/nclimate1633>.
- , K. E. Trenberth, and T. Qian, 2004: A global dataset of Palmer drought severity index for 1870–2002: Relationship with soil moisture and effects of surface warming. *J. Hydrometeorol.*, **5**, 1117–1130, <https://doi.org/10.1175/JHM-386.1>.
- D'Andrea, F., A. Provenzale, R. Vautard, and N. De Noblet-Decoudr, 2006: Hot and cool summers: Multiple equilibria of the continental water cycle. *Geophys. Res. Lett.*, **33**, L24807, <https://doi.org/10.1029/2006GL027972>.
- Duong, T., 2017: ks: Kernel smoothing. R package, version 1.10.7, <https://CRAN.R-project.org/package=ks>.
- Fischer, M., D. Kraus, M. Pfeuffer, and C. Czado, 2017: Stress testing German industry sectors: Results from a vine copula based quantile regression. arXiv, 12 pp., <https://arxiv.org/abs/1704.00953>.
- Ford, T. W., and S. M. Quiring, 2014: In situ soil moisture coupled with extreme temperatures: A study based on the Oklahoma Mesonet. *Geophys. Res. Lett.*, **41**, 4727–4734, <https://doi.org/10.1002/2014GL060949>.
- Genest, C., and A.-C. Favre, 2007: Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.*, **12**, 347–368, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2007\)12:4\(347\)](https://doi.org/10.1061/(ASCE)1084-0699(2007)12:4(347)).
- Gudmundsson, L., F. Rego, M. Rocha, and S. I. Seneviratne, 2014: Predicting above normal wildfire activity in southern Europe as a function of meteorological drought. *Environ. Res. Lett.*, **9**, 084008, <https://doi.org/10.1088/1748-9326/9/8/084008>.
- Hegerl, G. C., H. Hanlon, and C. Beierkuhnlein, 2011: Climate science: Elusive extremes. *Nat. Geosci.*, **4**, 142–143, <https://doi.org/10.1038/ngeo1090>.
- Hillier, J. K., N. Macdonald, G. C. Leckebusch, and A. Stavrinides, 2015: Interactions between apparently primary weather-driven hazards and their cost. *Environ. Res. Lett.*, **10**, 104003, <https://doi.org/10.1088/1748-9326/10/10/104003>.
- Hirschi, M., and Coauthors, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.*, **4**, 17–21, <https://doi.org/10.1038/ngeo1032>.
- Joe, H., 1997: *Multivariate Models and Multivariate Dependence Concepts*. CRC Press, 424 pp.
- Kraus, D., and C. Czado, 2017: D-vine copula based quantile regression. *Comput. Stat. Data Anal.*, **110**, 1–18, <https://doi.org/10.1016/j.csda.2016.12.009>.
- Kurowicka, D., and R. M. Cooke, 2005: Distribution-free continuous Bayesian belief. *Modern Statistical and Mathematical Methods in Reliability*, Series on Quality, Reliability and Engineering Statistics, Vol. 10, World Scientific, 309–322, https://doi.org/10.1142/9789812703378_0022.
- Leonard, M., and Coauthors, 2014: A compound event framework for understanding extreme impacts. *Wiley Interdiscip. Rev.: Climate Change*, **5**, 113–128, <https://doi.org/10.1002/wcc.252>.
- Lorenz, R., E. B. Jaeger, and S. I. Seneviratne, 2010: Persistence of heat waves and its link to soil moisture memory. *Geophys. Res. Lett.*, **37**, L09703, <https://doi.org/10.1029/2010GL042764>.
- Luo, L., D. Apps, S. Arcand, H. Xu, M. Pan, and M. Hoerling, 2017: Contribution of temperature and precipitation anomalies to the California drought during 2012–2015. *Geophys. Res. Lett.*, **44**, 3184–3192, <https://doi.org/10.1002/2016GL072027>.
- Martius, O., S. Pfahl, and C. Chevalier, 2016: A global quantification of compound precipitation and wind extremes. *Geophys. Res. Lett.*, **43**, 7709–7717, <https://doi.org/10.1002/2016GL070017>.
- Mazdiyasi, O., and A. AghaKouchak, 2015: Substantial increase in concurrent droughts and heatwaves in the United States. *Proc. Natl. Acad. Sci. USA*, **112**, 11 484–11 489, <https://doi.org/10.1073/pnas.1422945112>.
- Miralles, D. G., A. J. Teuling, C. C. van Heerwaarden, and J. Vil-Guerau de Arellano, 2014: Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nat. Geosci.*, **7**, 345–349, <https://doi.org/10.1038/ngeo2141>.
- Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**, D07S90, <https://doi.org/10.1029/2003JD003823>.
- Mo, K. C., and D. P. Lettenmaier, 2016: Precipitation deficit flash droughts over the United States. *J. Hydrometeorol.*, **17**, 1169–1184, <https://doi.org/10.1175/JHM-D-15-0158.1>.
- Mueller, B., and S. I. Seneviratne, 2012: Hot days induced by precipitation deficits at the global scale. *Proc. Natl. Acad. Sci. USA*, **109**, 12 398–12 403, <https://doi.org/10.1073/pnas.1204330109>.
- Noh, H., A. E. Ghouch, and T. Bouezmarni, 2013: Copula-based regression estimation and inference. *J. Amer. Stat. Assoc.*, **108**, 676–688, <https://doi.org/10.1080/01621459.2013.783842>.
- Palmer, W. C., 1965: Meteorological drought. U.S. Weather Bureau Research Paper 45, 58 pp., <http://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf>.
- Pham, M. T., H. Vernieuwe, B. De Baets, P. Willems, and N. Verhoest, 2016: Stochastic simulation of precipitation-consistent daily reference evapotranspiration using vine copulas. *Stochastic Environ. Res. Risk Assess.*, **30**, 2197–2214, <https://doi.org/10.1007/s00477-015-1181-7>.
- Porporato, A., and P. D'Odorico, 2004: Phase transitions driven by state-dependent Poisson noise. *Phys. Rev. Lett.*, **92**, 110601, <https://doi.org/10.1103/PhysRevLett.92.110601>.
- Rebel, K. T., R. A. M. de Jeu, P. Ciais, N. Viovy, S. L. Piao, G. Kiely, and A. J. Dolman, 2012: A global analysis of soil moisture derived from satellite observations and a land surface model. *Hydrol. Earth Syst. Sci.*, **16**, 833–847, <https://doi.org/10.5194/hess-16-833-2012>.
- Rebetez, M., O. Dupont, and M. Giroud, 2009: An analysis of the July 2006 heatwave extent in Europe compared to the record year of 2003. *Theor. Appl. Climatol.*, **95**, 1–7, <https://doi.org/10.1007/s00704-007-0370-9>.
- Ruffault, J., V. Moron, R. Trigo, and T. Curt, 2016: Objective identification of multiple large fire climatologies: An application

- to a Mediterranean ecosystem. *Environ. Res. Lett.*, **11**, 075006, <https://doi.org/10.1088/1748-9326/11/7/075006>.
- Salvadori, G., C. De Michele, N. T. Kottegoda, and R. Rosso, 2007: *Extremes in Nature: An Approach Using Copulas*. Water Science and Technology Library, Vol. 56. Springer, 292 pp.
- Schepsmeier, U., J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler, and T. Erhardt, 2017: VineCopula: Statistical inference of vine copulas. R Package, version 2.1.1, <http://CRAN.R-project.org/package=VineCopula>.
- Seneviratne, S. I., 2012: Climate science: Historical drought trends revisited. *Nature*, **491**, 338–339, <https://doi.org/10.1038/491338a>.
- , D. Lthi, M. Litschi, and C. Schr, 2006: Land–atmosphere coupling and climate change in Europe. *Nature*, **443**, 205–209, <https://doi.org/10.1038/nature05095>.
- , T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth Sci. Rev.*, **99**, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.
- , and Coauthors, 2012a: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field et al., Eds., Cambridge University Press, 109–230.
- , and Coauthors, 2012b: Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event. *Water Resour. Res.*, **48**, W06526, <https://doi.org/10.1029/2011WR011749>.
- Serinaldi, F., B. Bonaccorso, A. Cancelliere, and S. Grimaldi, 2009: Probabilistic characterization of drought properties through copulas. *Phys. Chem. Earth*, **34**, 596–605, <https://doi.org/10.1016/j.pce.2008.09.004>.
- Sheffield, J., E. F. Wood, and M. L. Roderick, 2012: Little change in global drought over the past 60 years. *Nature*, **491**, 435–438, <https://doi.org/10.1038/nature11575>.
- , and Coauthors, 2014: A drought monitoring and forecasting system for sub-Saharan African water resources and food security. *Bull. Amer. Meteor. Soc.*, **95**, 861–882, <https://doi.org/10.1175/BAMS-D-12-00124.1>.
- Sklar, M., 1959: Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris*, **8**, 229–231.
- Stagge, J. H., I. Kohn, L. M. Tallaksen, and K. Stahl, 2015: Modeling drought impact occurrence based on meteorological drought indices in Europe. *J. Hydrol.*, **530**, 37–50, <https://doi.org/10.1016/j.jhydrol.2015.09.039>.
- , D. G. Kingston, L. M. Tallaksen, and D. M. Hannah, 2017: Observed drought indices show increasing divergence across Europe. *Sci. Rep.*, **7**, 14045, <https://doi.org/10.1038/s41598-017-14283-2>.
- Teuling, A. J., and Coauthors, 2013: Evapotranspiration amplifies European summer drought. *Geophys. Res. Lett.*, **40**, 2071–2075, <https://doi.org/10.1002/grl.50495>.
- Törnros, T., and L. Menzel, 2014: Addressing drought conditions under current and future climates in the Jordan River region. *Hydrol. Earth Syst. Sci.*, **18**, 305–318, <https://doi.org/10.5194/hess-18-305-2014>.
- Trenberth, K. E., A. Dai, G. d. Schrier, P. D. Jones, J. Barichivich, K. R. Briffa, and J. Sheffield, 2014: Global warming and changes in drought. *Nat. Climate Change*, **4**, 17–22, <https://doi.org/10.1038/nclimate2067>.
- Tsakiris, G., and H. Vangelis, 2005: Establishing a drought index incorporating evapotranspiration. *Eur. Water*, **9**, 3–11, https://www.ewra.net/ew/pdf/EW_2005_9-10_01.pdf.
- Vautard, R., and Coauthors, 2007: Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit. *Geophys. Res. Lett.*, **34**, L07711, <https://doi.org/10.1029/2006GL028001>.
- Vicente-Serrano, S. M., S. Beguera, and J. I. López-Moreno, 2010: A multiscale drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J. Climate*, **23**, 1696–1718, <https://doi.org/10.1175/2009JCLI2909.1>.
- , and Coauthors, 2012: Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.*, **16**, 1–27, <https://doi.org/10.1175/2012EI000434.1>.
- , and Coauthors, 2014: Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environ. Res. Lett.*, **9**, 044001, <https://doi.org/10.1088/1748-9326/9/4/044001>.
- Wahl, T., S. Jain, J. Bender, S. D. Meyers, and M. E. Luther, 2015: Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nat. Climate Change*, **5**, 1093–1097, <https://doi.org/10.1038/nclimate2736>.
- Wand, M. P., and M. C. Jones, 1994: Multivariate plug-in bandwidth selection. *Comput. Stat.*, **9** (2), 97–116.
- Whan, K., J. Zscheischler, R. Orth, M. Shongwe, M. Rahimi, E. O. Asare, and S. I. Seneviratne, 2015: Impact of soil moisture on extreme maximum temperatures in Europe. *Wea. Climate Extremes*, **9**, 57–67, <https://doi.org/10.1016/j.wace.2015.05.001>.
- Zarch, M. A. A., B. Sivakumar, and A. Sharma, 2015: Droughts in a warming climate: A global assessment of standardized precipitation index (SPI) and reconnaissance drought index (RDI). *J. Hydrol.*, **526**, 183–195, <https://doi.org/10.1016/j.jhydrol.2014.09.071>.
- Zotarelli, L., M. D. Dukes, C. C. Romero, K. W. Migliaccio, and K. T. Morgan, 2010: Step by step calculation of the Penman–Monteith evapotranspiration (FAO-56 method). IFAS Publ. AE459, 10 pp., <http://edis.ifas.ufl.edu/ae459>.
- Zscheischler, J., and S. I. Seneviratne, 2017: Dependence of drivers affects risks associated with compound events. *Sci. Adv.*, **3**, e1700263, <https://doi.org/10.1126/sciadv.1700263>.
- , and Coauthors, 2014: Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on MsTMIP data. *Global Biogeochem. Cycles*, **28**, 585–600, <https://doi.org/10.1002/2014GB004826>.
- , R. Orth, and S. I. Seneviratne, 2017: Bivariate return periods of temperature and precipitation explain a large fraction of European crop yields. *Biogeosciences*, **14**, 3309–3320, <https://doi.org/10.5194/bg-2017-21>.